CHAPTER 5: BI sources and collection

'Just as the military draws upon a variety of intelligence sources, ranging from human intelligence through signals intelligence to radar intelligence, so will businesses need to develop the ability to 'broadcatch'. The task of the business intelligence centre will be to develop comparable all-source intelligence capability by a variety of methods' [4]

5.1 Introduction

One output of the BI requirements definition process is the identification of business intelligence gaps. The next stage of the business intelligence process, the collection stage, is aimed at addressing BI gaps.

According to Bernhardt (1993:171), collection 'involves the gathering of the raw information from which finished intelligence will be produced'. The purpose of the collection step of the intelligence process is not to collect as much as possible information regarding the KITs and KIQs as defined during the requirements definition process. Rather, the purpose is to selectively collect data, information and knowledge that will assist BI staff to address the identified BI gaps (as discussed in the previous chapter). Bernhardt (1993:173) emphasises the fact that the collection of information should be done 'as cost effectively, as efficiently and as quickly as possible'. In order to achieve this, the collection of information cannot be executed in an ad hoc manner, but needs to be properly planned and executed, using a collection process.

The aim of this chapter is to:

- Review existing literature on BI sources and collection processes

[4] B. Cronin (cited in Sigurdson and Tagerud, 2001:154)

- Discuss the BI collection processes used by SA banking institutions and to propose a BI collection process that could be applied by SA banking institutions.

Owing to the importance of the collection stage in the BI process, the collection of information is extensively discussed in literature on BI. However, in the literature the emphasis is often placed on the sources that could/should be utilised during the collection process. Typically, not much emphasis is placed on the actual steps to be followed during the collection process. To provide a practical step-by-step process for the collection of data/information/knowledge for BI assignments, this chapter considers the results of research done in SA banking institutions, as well as some BI collection practices used by international financial institutions.

5.2 BI sources

There is a myriad of BI sources that could be utilised during the collection stage of the BI process. In this regard it is important to gain an understanding of the various categories of BI sources in order to identify those sources that would be most appropriate to utilise when collecting information during a BI assignment.

5.2.1 Classification of BI sources

BI sources can be classified in a number of ways. Bernhardt (1993:171) distinguishes between two categories of data and information sources, namely primary sources and secondary sources. He notes that the primary sources are usually humans, whilst the secondary sources are 'open' sources. Kahaner (1998:53) also classifies information sources into primary and secondary sources. Gilad and Gilad (1988:58) distinguished between two types of information sources, namely field sources and published sources. Steele

(2000:128) refers to seven 'primary intelligence source types', which include imagery intelligence, human intelligence, signals intelligence, measurements and signatures intelligence, open-source intelligence, technical intelligence and counter-intelligence. Although Steele refers to these sources as 'intelligence sources', they are in fact sources that provide **data/information,** which is used as input for the generation of intelligence output.

5.2.2 BI source classification model

According to the definition of BI (as discussed in paragraph 2.7), BI sources include data sources, information sources and knowledge sources. Various classification methods exist to distinguish between BI sources. The following framework, based on the research conducted for this dissertation will be used to classify the various types of BI sources.

Primary Sources				Secondary Sources			
Internal		External		Internal		External	
Public	Non-Public	Public	Non-Public	Public	Non-Public	Public	Non-Public
Directed	Non- Directed	Directed	Non- Directed	Directed	Non- Directed	Directed	Non- Directed
Electronic	Non-Electronic	Electronic	Non-Electronic	Electronic	Non-Electronic	Electronic	Non-Electronic

Table 5.1: A framework for the classification of various types of BI sources

Each of these types of BI sources will be discussed in more detail below taking into consideration the nature and value of each of these types of BI sources.

5.2.3 Primary and secondary sources

5.2.3.1 Primary sources

(a) **Nature of primary sources**. Primary information sources refer to those sources that are the originators of the information. An example of this would be where a customer of a banking institution provides details about his/her financial needs to the banking institution. This information, in its original, unchanged format, is primary information. Kahaner (1998:53) refers to this as 'unadulterated facts directly from the source'. Continuing with the above example, when a staff member of this banking institution uses this primary information to compile a report on the customer's needs in which an opinion is expressed or other information elements are combined with the original information input, the staff member becomes a secondary source. Staff members observing the behaviour of customers in a banking hall are primary sources of information. Should there also be video footage of the banking taking place in the hall, the video material could also be classified as a primary source of information. Some publications may also be regarded as primary sources. Kahaner (1998:54) argues that speeches of senior management, annual reports and other company publications may be regarded as primary sources.

(b) **The value of primary BI sources**. According to Bernhardt (1993:173), 'it is primary resources that yield the most intelligence value'. Given the fact that the information is obtained from the origin of the information, primary source information, it is usually assumed to be more accurate than secondary source information. Kahaner (1998:54) agrees with this when he says, 'unless the source is deliberately lying, primary sources should be considered absolutely accurate'. Although it is often assumed that primary sources are more valuable for intelligence assignments than secondary sources, it should be noted that the

use of primary sources should be approached with caution as primary sources can provide information selectively. In the example of the banking customer providing the banking institution with details about his/her financial needs, the customer may forget or omit to tell the banking institution that he already had similar discussions with a competitor bank. This could lead to the institution's staff not accurately identifying that the customer is 'shopping around', or that they may lose this customer if they do not offer the same service and/or a better price than the competitor does. This does not diminish the importance of primary information sources for the BI projects of banking institutions. In 2000, research conducted by Ackerman and Wickens (2001:87) involving 33 financial services institutions found that 58 per cent of the institutions surveyed saw their customers as very important sources of information for intelligence assignments.

5.2.3.2 Secondary BI sources

(a) **The nature of secondary BI sources**. Kahaner (1998:54) defines secondary information sources as those sources that 'offer altered information'. Probably the best examples of secondary information sources are publications such as newspapers, magazines and analyst's reports, to name but a few. In the case of a publication, the writer of the publication could use material obtained during an interview (primary information) to compile an article that is edited prior to publication (secondary information). Edited television and radio broadcasts are also examples of secondary information sources.

(b) **The value of secondary BI sources**. In contrast to Bernhardt (see previous page), Kahaner (1998:55) is quick to point out that secondary sources are no less important or less accurate than primary resources. Kahaner (1998:55) maintains out that secondary information sources can

in some cases be more valuable than primary sources, especially where insightful opinions from the writer are shared in the publication. A typical example of this in a banking institution is where a report compiled by a market analyst (secondary information), providing expert opinions on the behaviour of customers, could prove more valuable to address a BI requirement than the original questionnaires (primary information) completed by customers during the market research.

5.2.3.3 A balance between primary and secondary BI sources

Since both primary and secondary information can immense value for banking institutions, there should be a balance between the development and utilisation of primary and secondary BI sources. An over-reliance on one type of BI source should be avoided. By comparing and integrating the information from both primary and secondary sources, BI analysts may in fact obtain a more comprehensive and balanced view of how best to deal with the BI gap.

5.2.4 Internal and external BI sources

5.2.4.1 Internal BI sources

(a) **The nature and value of internal BI sources**. Internal BI sources (primary and secondary) can play an important part in addressing intelligence gaps. It is not uncommon to find that the information/knowledge to address intelligence gaps is internally available, but that BI collectors do not know where to find it. There are many advantages to using internal BI sources. From a cost and/or risk perspective, it may be prudent to utilise internal BI sources before attempting to collect from external information sources. Especially when dealing with human sources, it is usually easier to obtain access to the internal sources than it is to make contact with and obtain access to external sources.

Unfortunately the opposite could hold true for some electronic sources, where it is not uncommon to find intelligence users and collectors gaining access to external databases, mainly because of the ease of access and navigation of such sources compared to internal databases. It should be noted that BI staff should not become too dependent on internal BI sources for collection purposes as these sources may provide information/knowledge that is biased towards organisational perspectives/policies.

(b) **Typical internal BI sources.** Several BI sources are utilised by banking institutions during the collection process. They include the following:

- **Employees**. Steele (2000:202) notes that 'the most important information resource is the employee'. He believes that every employee should form part of a bigger intelligence system and states that 'every employee must be a collector, producer and consumer of information and intelligence'. The research conducted by Ackerman and Wickens (2001:89) confirms the value of employees as collection sources for large international financial services institutions, especially those employees that have contact with banking customers (customer facing staff). Customer-care staff, relationship managers and sales staff are generally considered to be very important sources for BI purposes. One of the main advantages of using these employees in the collecting process is that they can collect the 'soft information', which is often difficult to collect and capture in the transactional processing systems of banking institutions. Examples of this include collecting information on customer attitudes, preferences, and information often volunteered by customers about their household, hobbies or future plans etc. In the case of SA banking institutions, 'relationship managers or personal bankers' are often utilised as sources for BI purposes, as they are in a position to collect 'soft' information during

customer contact. Typically sales staff members have knowledge of competitor products, marketing initiatives and the sales force, as they often have to compete against competing institutions when trying to conclude a sale.

- **Internal customers**. In the case of SA banking institutions it is not uncommon to find that employees are required to have accounts with their employer, often with preferential pricing structures. These employees are therefore also customers of the banking institution and make use of the products, services and channels of the institution. As such employees can provide primary information on their experiences of the institution's products and services. Employees can also provide information to increase the institution's understanding of customer attitudes, preferences and behaviour. Another benefit of utilising these employees as BI sources is that they can be approached with pertinent questions in a direct manner. Being internal to the organisation, the risk and cost of using these sources are relatively low when compared to using external sources. By asking an employee how and why he/she decided to acquire a specific product from the institution, BI staff members can obtain information regarding the decision-making processes of customers that have similar characteristics (e.g. age, income, qualifications, cultural background, lifestyle, etc.). Ackerman and Wickens (2001:89), referring to these BI sources as 'staff customers', point out that it was surprising that none of the 33 banking institutions that participated in their research viewed internal customers as an important BI source. The same applies to the SA banking institutions that participated in the research for the purpose of writing this dissertation, and appeared to emphasise the collection potential of staff members that have contact with customers, or can provide information on competitors.

- **Internal electronic sources.** Banking institutions need to maintain extensive internal electronic databases in order to support their operational processes. These internal databases can typically provide a wealth of customer data, product data, banking transactional data and channel data. In order to deal with the vast amounts of data stored in various electronic databases across banking institutions, it is common to find that data from various banking systems are integrated and stored in a central repository/database (a operational data store or a data warehouse). For banking institutions such a central database is a very important internal source of data/information, not only for BI purposes, but also for management reporting and analytics. It is especially for those BI assignments that relate to customers that such a repository becomes an important source. By sourcing customer data from all the operational systems and electronic customer touch points of a banking institution, BI collectors can obtain access to a very valuable internal source of customer data. Other typical internal databases that are utilised as BI sources include subject-oriented internal knowledge bases, internal library databases and internal intelligence databases.

5.2.4.2 External BI sources.

(a) **The nature and value of external BI sources**. External BI sources can also play an important part in addressing intelligence gaps. It is not uncommon to find that BI collectors often place too much emphasis on extracting information from internal sources due to the ease of access. There are many advantages to using external BI sources especially since these sources are not influenced by internal events or regulations of an organisation. The importance of external BI sources are confirmed by Peter

Drucker (1998:online) when he suggests that top management needs an information system that allows for 'the collection and organisation of outside focused information. All the data we have so far, including those provided by the new tools, focus inward'. The strong focus on internal data and information is also typical of banking institutions. Steele (2000:164) agrees with Drucker when he points out that 'the heart of the matter is not what you already know, or internal information, but rather what you don't know, or external information'. Bernhardt (1993:179) also emphasises the importance of external sources when he states that 'interviewing primary sources outside the organisation is the cornerstone of an effective CI collection process'.

(b) Typical external BI sources. Several external BI sources are typically utilised by banking institutions during the collection process. They include the following:

- **Business Publications**. All the SA banking institutions that participated in this research makes extensive use of business publications during the collection stage of the BI process. These publications range from the annual reports of competitors and corporate clients to press clippings and surveys published in business journals. In addition to this, all the SA banking institutions that participated in this research subscribe to the publications/reports published by industry research institutions.

- **Internet**. The importance of the Internet as a source of information during the collection stage of the BI process is confirmed by the research conducted for this dissertation. Despite issues around determining the accuracy of information published on the Internet, BI staff often use the Internet during as a starting point when looking for

external information. In this regard they would typically use the results of Internet search to identify other potential sources of information.

- **Commercial On-line Databases**. All the SA banking institutions that participated in this research makes extensive use of commercial online databases. These are typically used when BI staff require in-depth industry information or detailed company information that is not internally available.

5.2.4.3 Balance between internal and external BI sources

It should be noted that both internal and external sources could provide primary and secondary information (Bernhardt, 1993:173). Just as a balance has to be found between primary and secondary information sources, a balance between internal and external information sources is also essential. Whereas too much emphasis on internal sources may provide biased information, too much emphasis on external information could result in costly collection efforts that yield information that is not specific to the situation experienced within the banking institution.

5.2.5 Public and confidential BI sources

Another approach to categorise BI sources, is to distinguish between BI sources that are public or 'open', and those that are not in the public domain. It should be noted that both public and non-public BI sources could also be classified in terms of primary and secondary sources. These sources could also be internal and/or external to the banking institution.

5.2.5.1 Public domain 'open' sources

(a) The nature of public domain sources. In discussing these BI sources, Bernhardt (1993:177) refers to an infinite 'range of open sources – mainly published' that provides 'little in the way of real intelligence value'. Steele (2000:108) refers to open-source information as consisting of ' volumes of multi-media and multi-lingual information gathered for further processing and consideration'. Like Kahaner (1998: 59), Steele (2000:200) defines open sources as public domain information. Kahaner (1998:59) points out that in a regulatory business environment, organisations need to comply with regulations, which leaves 'a paper trail' of information. In addition to this 'regulatory' type of public domain information, there are various other types of public domain information that Kahaner describes as 'a vast sea of data that is open and available to anyone who seeks it'. Typical examples of open sources include, central and regional government departments, the media (printed and electronic), professional bodies, trade associations and the Internet. Kassler, (cited in Miller, 2000:113) confirms the importance of public sources, especially news resources, when she states that 'news sources have always served as essential tools in the intelligence research arsenal'. The ease with which news resources can be accessed makes them a popular BI source for collectors. Kassler mentions that news resources can be accessed in a number of ways, including 'print, commercial online archives, electronic alerting services and through the Internet'.

(b) The value of public domain sources. In comparing the qualitative value of 'open' sources with the value of primary resources, Bernhardt stipulates that the relative intelligence value of open sources is low. In contrast to Bernhardt's views, Steele (2000:128) emphasises the importance of 'open source information' (OSIF), which he argues is a critical input for the production of what

Steele terms 'open source intelligence'. Bernhardt seems to confuse the distinction made between primary and secondary sources with the difference between public and non-public sources of information. From the researcher's perspective, public or 'open sources' can also be primary sources. A typical example of an open source is the annual financial statements of a listed company, which contain primary information that is published in the public domain at the close of a financial year. The annual reports and financial statements of all the major SA banking institutions provide primary information and are available in the public domain.

5.2.5.2 Non-public domain and confidential sources

(a) **The nature and value of non-public domain sources**. It is not uncommon to find that information required to address a BI gap is not available from public domain sources. This does not imply that the information is private or confidential, but rather that it could be obtained if the correct source can be identified and is correctly approached. A typical banking example of this is where customer surveys are conducted to determine customer perceptions of a banking institution's service or market position. Typically such information would be extremely valuable to a banking institution but this is often not public domain information, unless it is published/made available to the public. Banking institutions often purchase this type of information from research companies. Kahaner (1998:80-87) refers to a number of methods used to gather non-public domain information and highlights some of the BI sources that are important in this regard. According to Kahaner, sales staff members are typical sources of non-public domain information. In their daily effort to sell a product/service, they not only have the opportunity to collect non-public-domain information, but they instinctively collect such information in order to assist them 'to make the sale'. Unfortunately salespeople rarely share their information on customers and

competitors unless they are provided with incentives to do so. One of the SA banking institutions that provided input for this dissertation has implemented an incentive programme to entice staff members to collect and share information of value for the BI programme.

(b) **The nature and value of confidential sources**. Unfortunately, BI is often associated with industrial espionage, which involves the use of illegal and unethical methods to obtain access to information that is confidential and classified. Steele (2000:128) claims that one of the reasons for this misconception is 'the conventional understanding of intelligence as information that is inherently classified'. In fact, confidential information can be collected for BI purposes in a legal and ethical manner. This usually requires that the collector obtain the consent of the BI source. A typical example of this is when confidential salary and personal asset information is collected by a bank, in order to determine a customer's ability to qualify for credit. It is of critical importance for banking institutions to obtain this confidential information as it assists the banking institution not only to determine the amount it would be lending to a customer but it also forms the basis for the banks financial risk management. Another example is where banking institutions collect information on their customers' credit card transactions, not only to provide their customers with banking institution statements, but also to understand customer lifestyles. This in turn can be used to segment the customer base and assist the banking institution to offer specific products/services to a customer.

5.2.6 Directed and non-directed sources

For the purposes of this dissertation the researcher would also like to classify BI sources as sources that can be directed, and those that cannot be directed. This

classification, often used by intelligence professionals, can assist collectors when they have to complete collection tasks under severe time constraints.

5.2.6.1 The nature and value of directed sources

Directed sources, which are often human BI sources, can be guided/directed by a collector to find very specific pieces of information. These sources are very valuable if they have access to the specific information required. A consultant or staff member being directed to obtain marketing material on a competing product from a competitor's branch office could be classified as a directed source. Electronic sources could also be directed to filter and search for specific pieces information as these become available in the public domain. Typically, directed sources should collect more relevant information/knowledge than sources that are non-directed. One major advantage of using these sources is that, if they are properly directed and they have the applicable access to information, relevant information could be obtained relatively quickly.

5.2.6.2 The nature and value of non-directed sources

These sources, i.e. those that cannot be directed, often provide irrelevant information/knowledge to the collector. Publications not commissioned by the banking institution or company websites are typical examples of non-directed sources in the public domain. BI collectors need to search through the information provided by these sources to find information that could be used to address the BI gap. The value of these BI sources are often determined by the accuracy and relevancy of the content.

5.2.7 Electronic and non-electronic sources

Another useful method for classifying BI sources is to distinguish between electronic and non-electronic sources. This classification can assist collectors in planning how to gain access to BI sources.

5.2.7.1 The nature and value of electronic sources. These sources abound and the focus is generally on remote access and the use of electronic search methods to obtain information. One of the challenges facing collectors when dealing with these sources, if they are external electronic sources, is to be able to verify information (especially if it is collected via the Internet). Electronic sources include websites of clients, competitors and research institutions, and commercial databases. Their value often lies with the ease of gaining access to these sources as well as the ease with which information can be electronically copied and manipulated.

5.2.7.2 The nature and value non-electronic sources

In the case of a human source, the collection of information would require making contact with the source. This provides an opportunity for two-way communication, as well as opportunities to verify the information obtained. Gilad and Gilad (1988:88) refer to these sources as field sources. The importance of using human sources or human networks for the collection of information is recognised by Cronin (cited in Sigurdson and Tagerud,1992:155), who refers to these networks as 'soft networks'. Cronin states that these networks could provide 'much potentially useful information, e.g., gossip, leads, tips, opinions, speculation, insights, and is transmitted through a variety of informal exchange mechanisms'. In addition to this Cronin notes that these networks form an essential component of the more formal and technology-based collection sources. It is important to note that Cronin refers to informal human source

networks and an informal exchange of information. Formal human source networks are common in banking institutions. Other non-electronic sources include printed media and other paper-based reports that could contain valuable information that may not be accessible via electronic means. Kassler (cited in Miller, 2001:98), referring to the printed media as 'the traditional source', confirms that printed sources remain an important source in BI assignments. Printed reference guides and business are often used as a starting point for the collection of business information.

5.2.8 Sources used by SA banking institutions

Based on the interviews conducted for this dissertation, the BI sources typically used by SA banking institutions can be depicted as follows:

Internal sources		External sources		
Non-electronic	Electronic	Human/ institutions	Non-electronic publications	Electronic
Publications in library Internal collection network BI champions and agents	Internal Intelligence System. Internal Knowledge system Data warehouse	Industry consultants. Intelligence consultants. Market research companies. Universities Government institutions	Industry reports Market reports	Commercial databases News-clipping services Internet

Table 5.2 Examples of sources typically used by SA banking institutions

From the above it is clear that these institutions make use of both internal and external sources, as well as human, electronic and non-electronic sources.

5.3. *Approaches to the collection process*

In the literature, two fundamental approaches to BI collection are identified. Kahaner (1998:56) refers to one approach that requires a specific BI requirement to be determined before the collection process is executed. This is similar to the reactive approach when dealing with BI requirements (as discussed in Chapter 4 paragraph 4.4.1). The second approach emphasises the need to collect information of interest on an ongoing basis. Such information is stored and updated so that it can be used when needed. This approach is similar to the proactive approach of dealing with BI requirements and is linked to the ability of BI staff to act on their own initiative. Kahaner (1998:56) suggests that both these approaches need to be followed.

According to Steele (2000:201), 'Just in time' collection and intelligence production, which he refers to as 'just in case collection and archiving', is far less expensive and far more useful than the proactive approach. Although the proactive approach suggested by Kahaner has advantages, it must be noted that it has some resource implications and may be practical only for very large corporations. Essentially, this proactive approach requires collectors to collect data, information and knowledge of interest without trying to address a specific intelligence gap. This is feasible if collectors have spare capacity and if an organisation has the necessary IT infrastructure to store and effectively retrieve this data/information/knowledge when it is needed.

It should be noted that the proactive approach to collection must not be confused with the ongoing collection tasks associated with intelligence requirements for current intelligence. From the researcher's perspective, BI programs in SA banking institutions have to produce BI products under both time and financial constraints, which suggests that Steele's approach of 'just in time' collection

may be the most practical approach to adopt. Although all three of the banking institutions that participated in this research obtain and store information flowing spontaneously from their internal collection networks, this flow of information is linked to BI requirements. In SA banking institutions there does not seem to be a dedicated BI effort to collect as much as possible information simply because it might become useful at a later stage. One of the banking institutions involved has built in a comprehensive 'filtering' process to prevent a situation where irrelevant information is pushed into the BI system.

5.4 Collection methods

The methods applied to obtain information/knowledge from BI sources would typically vary, depending on the type of source to be used in the collection process and the associated time and budget constraints. Professional BI collectors place emphasis on identifying and using the most appropriate method in relation to a specific source, in order to ensure that the collection effort is as effective as possible. Electronic searches are generally considered to be the most appropriate method to use when collecting information from a searchable electronic source. When collecting information from human sources, conversations and interviews are typical methods to be used, but depending on the individual to be interviewed, a collector may decide to use a specific interviewing technique that would be the most appropriate. When collecting information on new products, and appropriate method to use would be observing and/or purchasing an example of such a banking product.

It is important to note that the methods used during the collection process should be ethical and legal. Moving beyond these boundaries could be considered to be industrial espionage. Within a banking institution, all BI staff members and those involved with BI should be able to distinguish between collection methods that are considered to be appropriate and those that should be avoided.

The Royal Bank of Canada (2001) provides a basic model of collection methods that could be used to obtain primary information on competitors and customers. This banking institution emphasises that the specific focus should be on the collection of competitor and customer information, and suggest the following methods:

Methods to collect primary information on competitors	Methods to collect primary information on customers
Store visits	Phone interviews
Test goods	Face-to-face meetings
Obtain competitor reports	Mail surveys
	Focus groups

Table 5.3 Collection methods (Royal Bank of Canada,2001)

Although the above table does not provide an extensive list of collection methods, it provides an indication of the methods typically considered by a banking institution to be of use for BI purposes. The importance of events for collection purposes also needs to be briefly discussed.

Events, whether they are internal or external company/industry events, provide an opportunity to collect information. Although an as such cannot be classified as a BI source, there are methods specifically developed to collect information during events. During internal events, banking institutions can determine that they need to invite specific guests in order to create an opportunity for BI collectors to make contact with these guests. Typically collectors would compile contact reports after having met with these guests. External events, such as industry-specific events or trade shows, provide an opportunity to collect non-public-domain information by making contact with potential customers,

competitors and/or potential suppliers. Some typical methods used during events include observing the products/services being offered by competitors and observing potential customers visiting the exhibitions of competitors.

5.5 *Collectors of information*

Staff members who perform the task of collecting information from BI sources, can be referred to as collectors. When planning and structuring a collection function or team, any one of several approaches could be followed. One approach is to use a select group of BI staff members to execute the collection process. Another approach is to establish an internal collection network that also includes staff members who are not dedicated BI specialists. Gilad and Gilad (1988:55) prefer the latter approach and point out that the internal collection network consists of staff members that are not the sources of information, but are 'those individuals who access sources of information'.

Even though an organisation may have an internal collection network, the services of external collectors may be required. In this regard, intelligence research companies, market research organisations and consultants can be used as collectors. Gilad and Gilad (1988:87) point out that 'both internal and external collection are useful, each with its own limitations'. One of the reasons for making use of external collectors is to protect the identity of the company that requires the information to be collected. Another reason for making use of external collectors could relate to the limitations of the internal collection network. It could be that internal collectors may not have any available sources to exploit in a particular market, whilst the external collector may have already established information sources in that particular market. External collectors with the required access to sources can prove valuable for the collection process especially when there are time constraints to deal with on an intelligence assignment. Gilad and Gilad (1988:87) cautions against placing too much

emphasis on external collectors and suggest that external collectors should supplement the internal collection network.

Although the concept of an internal collection network, as suggested by Gilad and Gilad, has many advantages from the researcher's perspective, some collection tasks require specialist collectors. Not all employees are necessarily skilled or have the personality traits of an effective interviewer or information scientist. The researcher believes that an effective collection structure should consist of specialist BI collectors, supported by both an internal and an external collection network.

Research conducted amongst SA banking institutions indicated that all the participating banking institutions emphasised the need to establish and maintain internal collection networks. All of the banking institutions also made use of external collectors to support the internal network. Only one of the participating banking institutions appeared to have established an external collection network. None of the participating banking institutions seemed to have established dedicated BI collection teams. Typically, BI staff members were involved in both collection and analysis activities.

5.6 A process for the collection of information

5.6.1 Different perspectives on collection processes

When Bernhardt (1993:171) refers to collection and processing as 'respectively the second and third steps in the intelligence cycle', it is clear that he refers to the stages of the intelligence process as designed by the CIA. The processing stage of this specific intelligence process is aimed at converting/transforming the data/information collected into a format that is suitable for the analysts to deal with during the next stage of the process (analysis stage). Due to the fact that

the CIA intelligence collection effort includes the collection of vast amounts of signals, radar and imaging information, a specific stage was required to convert data from various formats and often from various languages into a format that could be used by analysts. Due to the nature of information collection for BI purposes in the banking environment, a specific processing stage is not required. Should the data/information collected require conversion or translation, e.g. from a foreign language, this would typically be done as part of the collection process.

In the relevant literature, various approaches to collecting information for BI purposes are identified, including the traditional approaches and the 'diamond paradigm'.

5.6.1.1 Traditional approaches

Bernhardt (1993:172) suggests a collection process consisting of four steps. These steps should be executed in the following sequence:

- Determine the research design.
- Determine the data and information collection methods.
- Design the researcher's collection forms and, if possible, agree on these with the intelligence consumer.
- Collect the data and information.

The researcher has some concerns regarding this process. By referring to a research design, this process seems to include some of the steps of the requirements definition process. The KITs and KIQs identified in the BI requirements process provide a framework or 'research' design for the BI assignment. The focus of the collection stage of the intelligence process should be on the planning and collection of data/information/knowledge. This process also makes no reference to the identification of the most appropriate sources to

utilise during the collection process. As mentioned in the paragraph 5.4, the type of source selected would determine the most appropriate method to use during the collection process.

The collection process suggested by Gilad and Gilad (1988:56) can be summarised as follows:

- Establishment of an internal collection network
- The identification of sources of information
- Obtaining access to sources of information
- Collection of the relevant data/information
- Report information according to pre-established communication procedures

In this process, the identification of sources (which is lacking in Bernhardt's process) is a prominent step, which is concluded before any method is applied to obtain access to sources. The reporting of information according to a predefined communication procedure ensures that information/knowledge collected during the process is captured in a standard format, which is important for later retrieval. The only point of concern regarding this process is that it places too much emphasis on the internal collection network (also refer to paragraph 5.5).

A generic collection process, which conforms to the traditional or 'linear' approach to collection, can be depicted as follows:

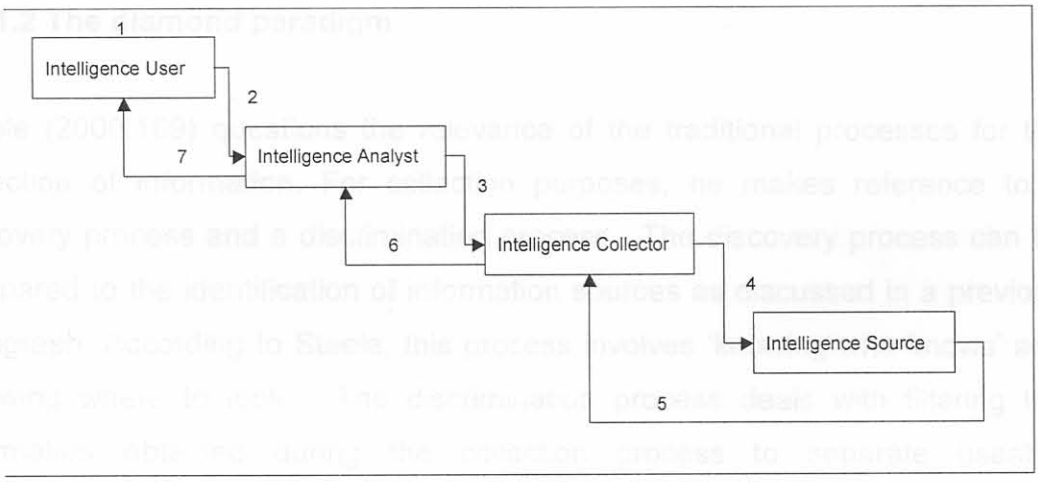

Figure 5.1: The linear collection approach

The first step (1) requires the intelligence user to determine his/her business problem and then to communicate this to a BI analyst (2), who will assist with defining BI requirements. The analyst will task the most suitable BI collector/s (3) who have knowledge of and access to the information sources. The collectors (4) will gain access to the sources and data/information/knowledge will be obtained from the source (5). The information will be verified, or converted if required, and given to the analyst (6). The analyst will then evaluate the information in terms of relevancy and accuracy. On completion of this, the analyst will perform a process to generate intelligence output before providing this to the intelligence user (7). Should the user request receipt of the information without any analysis having been performed on it, the information will pass from the analyst to the intelligence user. According to Steele, this process is too slow and not workable in a fast-moving business environment.

5.6.1.2 The diamond paradigm

Steele (2000:109) questions the relevance of the traditional processes for the collection of information. For collection purposes, he makes reference to a discovery process and a discrimination process. The discovery process can be compared to the identification of information sources as discussed in a previous paragraph. According to Steele, this process involves 'knowing who knows' and 'knowing where to look'. The discrimination process deals with filtering the information obtained during the collection process to separate useable information from information that would not be useful for the intelligence process. Probably his main departure from the traditional approach relates to the interaction between various role players during the collection process.

Steele (2000:201) promotes an approach that he refers to as the diamond paradigm. In this approach the intelligence user, the BI collector, the BI analyst and the source communicate directly with one another and information flows freely between these role players.

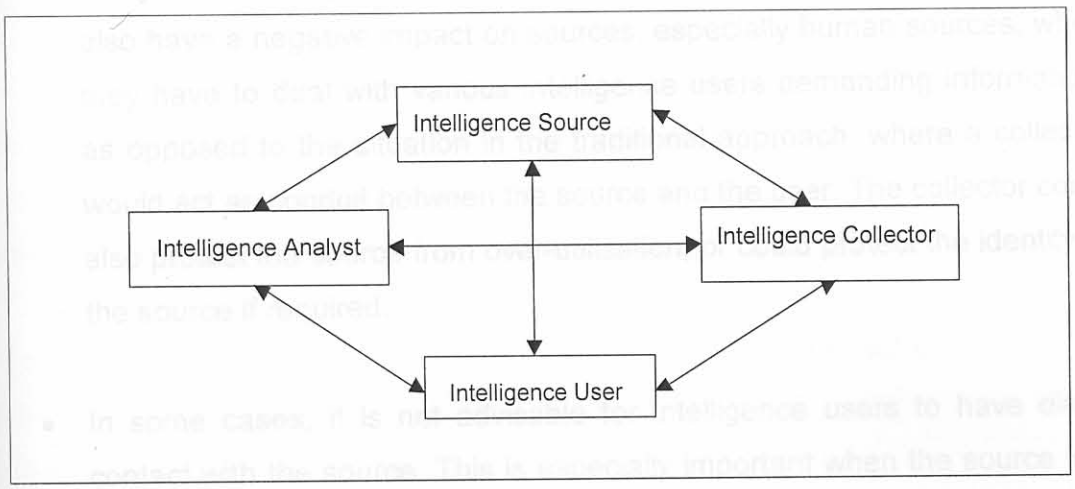

Figure 5.2: The Diamond Paradigm

Although Steele's approach is not without merit, especially when intelligence assignments have to be conducted under severe time constraints, a number of negative implications are associated with it. These include the following:

- It is possible that **information** that has not been verified or evaluated could flow directly between source and intelligence users due to time constraints. This could lead to a situation where the intelligence user uses 'raw information' as the basis for decision-making and taking action. The onus is on the intelligence user to evaluate the raw information, or refer it to the analyst for review.

- From an intelligence-management perspective, the diamond paradigm provides some challenges. One challenge is to ensure that the intelligence user is linked to the most appropriate source and that the most appropriate method of collection is used. In the diamond paradigm the user can make contact directly with the source without the collector or analyst necessarily having knowledge of such contact. This situation could also have a negative impact on sources, especially human sources, when they have to deal with various intelligence users demanding information, as opposed to the situation in the traditional approach, where a collector would act as conduit between the source and the user. The collector could also protect the source from over-utilisation, or could protect the identity of the source if required.

- In some cases, it is not advisable for intelligence users to have direct contact with the source. This is especially important when the source is a member of a competitor organisation, or where the possibility exists that the intelligence user's ultimate intentions would become apparent when

dealing directly with sources.

- Finally, it should be noted that unless users are trained in collection methods and are proficient in them, involving them directly in the collection could be less effective than following the linear process. A typical example of this would be where an intelligence user obtains access to an on-line database, but does not know how to develop a search strategy and is unable to use the method that would be best for searching the particular database. By giving this collection task to a qualified information scientist, the electronic search could potentially be completed with more accuracy and in less time.

During the research conducted n SA banking institutions, it became evident that not one of the participating banking institutions have implemented the diamond paradigm as described by Steele. In all the participating banks, there is an informal/free flow of information between analysts, collectors and sources. In two of the banks there are BI champions/agents for each business unit, who are also responsible to act as collectors and conduits for collection tasks. In none of these banks a situation exists where intelligence users actively engage with BI sources as part of the BI collection process without the involvement of BI staff. One of the banking institutions has adopted a more formal 'linear' approach to collection. In order to prevent their process from becoming slow and inefficient, an IT system was designed and developed to facilitate the collection process.

A seven-step process for BI collection, based on the traditional approach and Steele's diamond paradigm, as well as the researcher's practical consulting experience of BI collection, is suggested.

5.6.2 A seven-step process for BI collection

In the light of Steele's criticism of the traditional approach towards collection in the intelligence environment and the need for a more dynamic process, the following process could be considered for the collection stage of the BI process. It should be noted that this process is designed for BI programmes with established internal and external source networks, and access to the services of specialist BI collectors.

This seven-step process is more pragmatic than the typical linear approach in that it allows for the involvement of intelligence users during different stages of the process. It emphasises the importance of understanding the sources available and then planning the collection effort before collecting data, information and knowledge. In this respect, the process remains true to the fundamentals of collection as practised in the intelligence profession.

The process is triggered by the 'intelligence gap', as described in Chapter 5 of this dissertation, and could be depicted as follows:

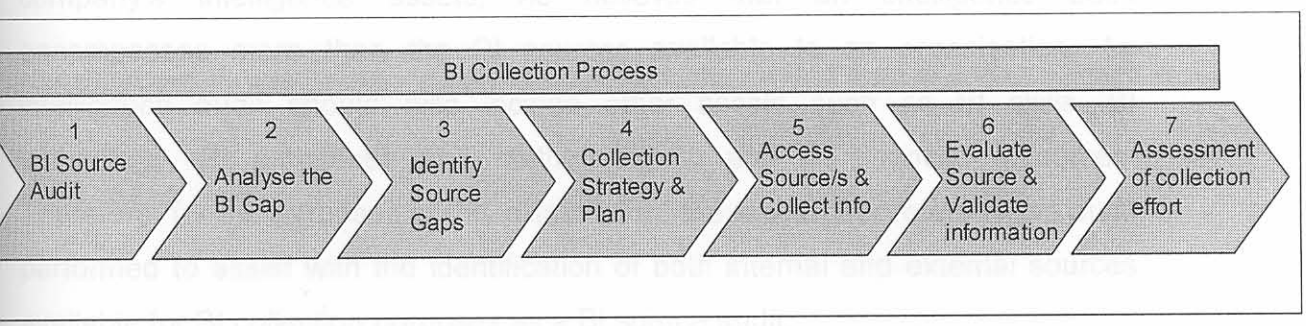

Figure 5.3: A seven-step BI collection process

Step 1: BI source audit. Knowing which sources are available to BI collectors is a critical component of this collection process. The purpose of this audit is mainly

to identify the existing data, information and knowledge sources, and networks already available (both internal and external) to BI collectors. Secondly, the purpose is also to determine the extent to which these existing sources and networks are being utilized.

Generally speaking, there is a wealth of BI information sources at the disposal of SA banking Institutions. One of the banking institutions that participated in this research confirmed that it had completed such an audit as part of the collection process. The importance of a BI sources audit is confirmed by Bernhardt's (1993:173) claim that most of the information required for competitive intelligence purposes is to be found within organisations, but that 'the challenge lies in finding it'. Fuld (1988:64) suggests that, in order to identify an organisation's information sources, an intelligence audit should be conducted. He defines an intelligence audit as an 'inventory of your **company's intelligence assets'**. To this he adds: 'These assets include private competitor files, individually constructed databases, market studies purchased outside the library, as well as names of industry and competitor experts within your company.' Although the researcher agrees with Fuld's definition that an intelligence audit is an inventory of a company's intelligence assets, he believes that an intelligence audit encompasses more than the BI sources available to an organisation. An intelligence audit should also include other assets such as BI skills, BI technology, BI processes and methodologies, and BI financial resources. Therefore, for the purposes of this dissertation, the researcher refers to the audit performed to assist with the identification of both internal and external sources available for BI collection purposes as a BI source audit.

A BI source audit may identify a surprising number of valuable internal and external information and knowledge sources that are available and accessible to BI collectors. In addition to the identification of sources that can be used for BI

purposes, such an audit could also indicate reliance on only a small number of sources whilst a number of others are under-utilised. It is important to note that BI source audits need to be conducted periodically, as new sources may appear and others may not be available for further use.

The BI source audit can also assist in identifying areas of source duplication. In this regard it is not uncommon to find a number of departments within a banking institution subscribing to the same publication.

The identification of the existing data/information/knowledge sources is but one part of the BI source audit. Once identified, the sources must be classified as per the classifications of Table 5.1. In addition to this, the area of expertise of each source should also be identified. This should indicate which sources could be utilised to obtain expert input on specific topics. Sources should also be classified in terms of the accuracy of information provided and their reliability. Steele (2000:109) refers to this as the careful discrimination between good and bad sources, relevant and irrelevant sources, and finally between cost-effective and cost-prohibitive sources'.

It is also important that as part of the audit, sources should be classified in terms of their access to information. Some human sources may have access to the executive management and strategic planning information of competitors, while for others access may be restricted to operational activities. The same classification could be applied to electronic sources, as some databases may only provide access to specific topic-related information.

An example of a matrix that could be used during the BI source audit is depicted below:

BI source	Source type (Primary/secondary) (Internal/external) (Electronic/non-electronic) (Directed/non-directed)	Reliability of source	Accuracy of information	Access to information (Topics) (Types) (Levels in organisation)	Constraints (Costs) (Time)
Source 1					
Source 2					

Table 5.4 BI source audit

This matrix must be updated at regular intervals by obtaining input from Step 6 (evaluation of sources and validate information) and Step 7 (assessment of collection effort) of the collection process.

In addition to the above, it is important to understand the 'idiosyncrasies' of BI sources. Kahaner (1998:65) emphasises the importance being thoroughly familiar with BI sources as 'some are more complete, some oversimplify and some have biases'. A typical example of this is where BI collectors, after using a publication for some time, can identify those writers/reporters that are more knowledgeable and whose articles are more accurate that those written by other reporters covering the same topic or contributing to the same publication.

Step 2: Analysis of the BI gap. Typically, this step would require BI analysts, collectors and intelligence users to discuss the information gaps and reach agreement on how the collection task should be approached. During this step it is also important to confirm the roles and responsibilities involved in the

collection process. It could be decided that intelligence users need to form part of the collection team, and that they need to have access to the sources of information due to time constraints. This is similar to Steele's idea of having the intelligence user involved in the process, but this involvement will be agreed 'up front' and the collector/s will be part of the process. Another important task associated with this step is the reviewing of the KIQs that were formulated as part of the BI requirements definition process. In their report on their research, Ackerman and Wickens (2001:81) make specific reference to the fact that 'formulated KIQs may not contain enough detail for collection staff to plan and execute the collection task'. When KIQ's are analysed at a lower level, collectors are able to start identifying the most appropriate sources and techniques to use during the collection effort. A typical banking KIQ could relate to determining the reasons why certain groups of customers move their banking accounts from a competitor to the banking institution conducting the BI assignment. In this case, posing the question in this 'KIQ' format to a particular group of customers could result in some customers providing answers and others deciding not to share their motives. Responses may be more forthcoming if this KIQ could be changed to a collection question, e.g. 'Would you care to tell us what attracted you to our banking institution?'

Step 3: Identify BI source gaps. On completion of the above step, the BI collector should have a clear view of 'what' data, information and knowledge need to be collected. The next step is to identify which of the internal and/or external sources would be the most appropriate for obtaining the required information. In this regard the importance of having completed BI source audits needs to be emphasised. Unless the collector has a clear picture of all the BI sources at his/her disposal, chances are that he/she may utilise a source that is not suitable or, even worse, identify a source gap. Source gaps develop when the information required cannot be obtained from any of the existing data,

information and knowledge sources. Typically, BI source gaps are dealt with by developing new sources and expanding the collection network.

Step 4: Compile the collection plan. In order to prevent the collection effort from turning into an ad hoc process, where collectors, analysts and intelligence users collect information from sources without any form of coordination, the collection effort must be properly planned. This is especially important when customers are used as BI sources, and when information on competitors is being collected. When direct contact with a customer is deemed to be an appropriate method of obtaining information, the banking institution would not want several staff members contact the same customer and pose the same questions. The same applies to the collection of competitor information. If a number of staff members phone a branch of a competing banking institution and ask the same questions, the competitor banking institution might realise that a BI collecting process is being conducted. Similarly, a collection effort focussed on electronic searches should also be planned. In most cases the most appropriate database, or databases, and the query method or language to be used need to be identified. Access may require subscription if it is an external database. Having a specific set of questions, searching in more than one language, or even making use of acronyms or different spellings of key words used in the search, could form part of the search strategy when exploiting a text-based database. Determining a collection strategy and planning the collection effort are essential to ensure the optimal utilisation of BI collectors and information/knowledge sources.

- **The collection plan.** In order to execute their collection tasks, collectors need an operational collection plan. The purpose of this collection plan is to determine which intelligence gap is to be dealt with by which collector, and to indicate the sources that will be utilised and the collection methods

that will be applied. Generally, the first step is the listing of all the BI requirements, KITs and KIQs, and the intelligence gaps as defined in the BI requirements definition process. All the specific collection questions related to these intelligence gaps, as identified during the analysis step of the collection process, should also be listed. Once this list is complete, all the potential sources of information/knowledge that could provide the required information are listed per intelligence gap. In order to do this effectively, the collectors need to know the capabilities of all the information sources at their disposal. The next step is to identify the most appropriate source/s to provide answers to the formulated questions. Once this has been done, the most appropriate method for collecting the information from the specific source is determined. By doing this an initial collection plan is established, which could be depicted as follows.

Intelligence requirement and gap	KITs	KIQs	Specific collection questions	Available sources	Appropriate source/s	Method for collection	Collector

Table 5.5 The collection plan

The collection plan could be expanded to include the prioritisation of the intelligence requirements and a discussion of the constraints to which the completion of the task will be subjected. Of specific importance is the time and budget constraints associated with the collection task. In some cases, BI managers would add target dates to the plan and monitor the status of the collection effort using this plan.

One of the banking institutions that participated in this research, places a lot of emphasis on having an updated collection plan, similar to example depicted

above. This institution has an overall collection plan for the bank, which integrates all the BI collection plans of the various business units. The business unit collection plans typically contain the BI sources available to that business unit. This allows staff at the corporate headquarters level to have an overview of all the collection tasks in the banking institution and limits duplication of effort.

Step 5: Obtain access to sources and collect information. On completion of a collection plan, the collection tasks can be initiated. This typically starts with the collector having to establish contact with and gain access to the source. The next step is to pose the collection questions/query and to obtain the relevant data, information and knowledge from the source. When dealing with a human source, and depending on the most appropriate method, the collector may need to make an appointment and conduct an interview. Sometimes a brief telephonic interview will suffice. When dealing with electronic sources, especially those that require subscription, the collector may be required to register as a user before he/she will be able to gain access to sources and make queries. An integral part of this step is to capture and store all the data/information/knowledge obtained. In this regard, a specific database may be required where this 'raw' input from sources can be stored until the collectors can perform the validation step. One of the banking institutions that participated in this research uses a standard format for capturing all information collected in a collection report format.

Step 6: Evaluate source/s and validate information. Once the 'new' or 'raw' data, information and/or knowledge has been collected, it should be validated and, if required, converted into the format that is required by the analyst/intelligence user. It is also important to evaluate the source/s used during the collection process. A model developed for the validation of collected information/knowledge and evaluation of sources (Ackerman and Wickens:

2001:84) could be adapted and used by collectors during this step of the collection process.

A derivative of this model is depicted below:

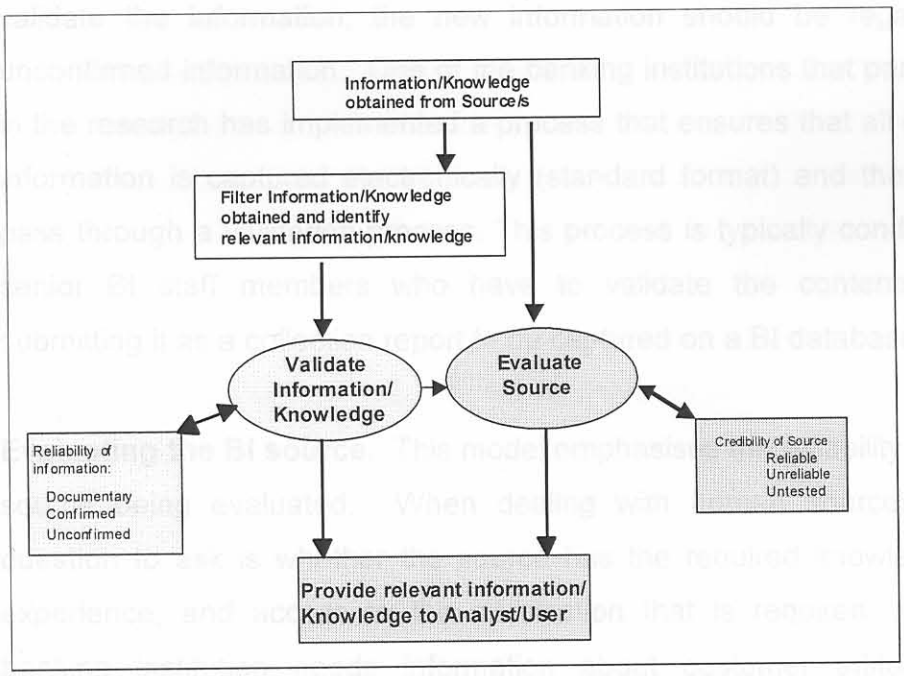

Figure 5.4: A model for evaluation of sources (Ackerman and Wickens, 2001:84)

- **Validation of information/knowledge**. In this model collectors filter all the bits of collected data/information/knowledge in terms of their relevance in addressing the collection questions and intelligence gap as per the collection plan. Only the relevant data/information/knowledge is then validated in terms of the reliability of the information. To determine the reliability of the 'new' information/knowledge, it is usually compared with the existing data/information/knowledge/intelligence within the organisation, and by comparing/testing it against other sources that also provided information/knowledge on the same topic. In cases where other

independent sources confirm the information, the information can be classified as confirmed. Kahaner (1998:56) also emphasises the importance of validating information obtained from both primary and secondary sources. In cases where collectors/analysts are unable to validate the information, the new information should be regarded as unconfirmed information. One of the banking institutions that participated in the research has implemented a process that ensures that all collected information is captured electronically (standard format) and then has to pass through a validation process. This process is typically conducted by senior BI staff members who have to validate the contents before submitting it as a collection report to be captured on a BI database.

- **Evaluating the BI source.** This model emphasises the reliability of the BI source being evaluated. When dealing with human sources, a key question to ask is whether the source has the required knowledge and experience, and access to the information that is required. When a banking institution needs information about customer attitudes and preferences, the customers themselves are better able to provide reliable information than an external database or the media. In this model sources are classified as reliable, unreliable and untested. By keeping records of data/information/knowledge collected from these sources and the reliability of the data/information/knowledge they provided in the past, the historical performance and reliability of each source could be established. This can assist collectors to evaluate sources efficiently.

- **Providing validated information**. On completion of the validation, analysis and conversion tasks, the collector needs to make the data/information/knowledge available to analysts/intelligence users. This could be done through verbal communication, or by using the standard-

format 'collection reports" that are used in some banking institutions. One of the banking institutions that participated in this research has implemented a process to facilitate the capturing of collection reports in a standardised format on the corporate intranet.

Step 7: Assessment of the collection effort. In order to assess the efficiency and effectiveness of the collection process, Ackerman and Wickens (2001:85) suggest that a workflow system should be considered to monitor the progress of all collection tasks and to update the collection plan. The collection effort needs to be assessed in terms of addressing the BI gaps. In this regard, feedback would typically be obtained from analysts and intelligence users regarding the accuracy and relevance of the data/information/knowledge collected. In addition to this, the collection tasks also need to be assessed in terms of the extent to which they succeeded in providing the required information within time and budget constraints.

5.7 Conclusion

Vast numbers of data, information and knowledge sources are available to SA banking institutions. In order to identify the most appropriate BI source and then to obtain access to such a source, the BI collector should have a list of all the BI sources at his/her disposal. In this regard the importance of the BI source audit and the development of a collection plan cannot be stressed enough.

The research conducted in SA banking institutions confirms that many BI sources are utilised, but that the institutions involved may not be exploiting all their BI sources optimally, or do not have a list of all the BI sources available. Only one of the institutions confirmed that they had completed a BI source audit. Although SA banking institutions focus on developing internal collection networks, the value of using internal customers as part of these networks

appears to be underrated. Also, the development of formal external collection networks seems to be limited.

It is evident that there is a realisation of the need for a formal collection process, as all the institutions involved had already implemented such processes. The process implemented by one of the banks includes several of the steps of the seven-step process recommended in this chapter.

Ultimately, the collection stage of the BI process is aimed at addressing BI gaps, and in order to do so effectively and efficiently, BI collectors must have sound knowledge of available BI sources. This 'knowledge' goes beyond being able to identify and classify each BI source, and refers the ability to decide on the most appropriate source, based on the source's accessibility and credibility, regardless of whether it is a primary or secondary source.